

Title: Multiple sequence alignment using genetic algorithms

Author: Zdeněk Pátek

Department: Department of Software and Computer Science Education

Supervisor: RNDr. František Mráz, CSc.

Abstract:

The thesis addresses the problem of multiple sequence alignment (MSA). It contains the specification of the proposed method MSAMS that allows to find motifs in biological sequences, to split sequences to blocks using the motifs, to solve MSA on the blocks and finally to assemble the global alignment from the aligned blocks and motifs. Motif search and MSA are both solved using genetic algorithms. The thesis describes the implementation of the method, configuration of its settings, benchmarking on the BALiBASE database and comparison to the ClustalW program. Experimental results showed that MSAMS can discover better alignments than ClustalW.

Keywords: multiple sequence alignment, motif finding, genetic algorithms, ClustalW